Raloxifene for women with Alzheimer disease

A randomized controlled pilot trial

Victor W. Henderson, MD, MS
Tom Ala, MD
Kristin L. Sainani, PhD
Allan L. Bernstein, MD
B. Sue Stephenson, MD
Allyson C. Rosen, PhD
Martin R. Farlow, MD

Correspondence to Dr. Henderson: vhenderson@stanford.edu

ABSTRACT

Objective: To determine whether raloxifene, a selective estrogen receptor modulator, improves cognitive function compared with placebo in women with Alzheimer disease (AD) and to provide an estimate of cognitive effect.

Methods: This pilot study was conducted as a randomized, double-blind, placebo-controlled trial, with a planned treatment of 12 months. Women with late-onset AD of mild to moderate severity were randomly allocated to high-dose (120 mg) oral raloxifene or identical placebo provided once daily. The primary outcome compared between treatment groups at 12 months was change in the Alzheimer's Disease Assessment Scale, cognitive subscale (ADAS-cog).

Results: Forty-two women randomized to raloxifene or placebo were included in intent-to-treat analyses (mean age 76 years, range 68–84), and 39 women contributed 12-month outcomes. ADAS-cog change scores at 12 months did not differ significantly between treatment groups (standardized difference 0.03, 95% confidence interval −0.39 to 0.44, 2-tailed $p = 0.89$). Raloxifene and placebo groups did not differ significantly on secondary analyses of dementia rating, activities of daily living, behavior, or a global cognition composite score. Caregiver burden and caregiver distress were similar in both groups.

Conclusions: Results on the primary outcome showed no cognitive benefits in the raloxifene-treated group.

Classification of evidence: This study provides Class I evidence that for women with AD, raloxifene does not have a significant cognitive effect. The study lacked the precision to exclude a small effect. *Neurology*® 2015;85:1937–1944

GLOSSARY

AD = Alzheimer disease; **ADAS-cog** = Alzheimer's Disease Assessment Scale, cognitive subscale; **CI** = confidence interval; **SERM** = selective estrogen receptor modulator.

The burden of Alzheimer disease (AD) falls heavily on women. By virtue of greater longevity, more women than men survive to an older age when risk is greater. AD pathology is more likely to be expressed as dementia in women than men,[1] and cognitive symptoms appear to be more severe.[2] Estrogens have attracted interest as potential treatment for women with AD, but relatively small therapeutic trials have generally failed to confirm efficacy.[3] A number of compounds that lack the 4-ring cyclopentanophenanthrene structure characteristic of steroid hormones interact with estrogen receptors or exhibit estrogen-like properties. The selective estrogen receptor modulators (SERMs) have estrogenic effects in some tissues but antiestrogenic effects or no estrogenic effect in other tissues. Raloxifene, an oral SERM, is approved for treatment of osteoporosis in postmenopausal women. In the Multiple Outcomes of Raloxifene Evaluation trial, high-dose (120 mg/d) raloxifene reduced the risk of mild cognitive impairment (relative risk 0.67, 95% confidence interval [CI] 0.46–0.98) and led to a nonsignificant reduction in AD risk (0.51, 95% CI 0.21–1.21).[4] Studies of raloxifene effects on cognition in women

Supplemental data at Neurology.org

From the Departments of Health Research and Policy (Epidemiology) (V.W.H., K.L.S.), Neurology and Neurological Sciences (V.W.H.), and Psychiatry and Behavioral Sciences (A.C.R.), Stanford University, CA; Department of Neurology (T.A.), Southern Illinois University School of Medicine, Springfield; Neurology (A.L.B.) and Family Medicine (B.S.S.) Departments, Kaiser Permanente Santa Rosa Medical Center, CA; Palo Alto Veterans Affairs Health Care System (A.C.R.), CA; and Department of Neurology (M.R.F.), Indiana University School of Medicine, Indianapolis. A.L.B. is currently affiliated with Sonoma West Medical Center, Sebastopol, CA.

Go to Neurology.org for full disclosures. Funding information and disclosures deemed relevant by the authors, if any, are provided at the end of the article.

without dementia provide inconsistent results,[5–10] but some suggest benefit on a subset of neuropsychological measures.

Raloxifene has not been evaluated as treatment for AD. In consideration of large-scale efficacy trials, we undertook a pilot trial in postmenopausal women with late-onset AD of mild to moderate severity to determine whether raloxifene improves cognitive function and to provide an estimate of the treatment effect.

METHODS

Design and setting. This pilot study was conducted as a randomized, double-blind, placebo-controlled trial, with enrollment between September 2006 and December 2009 and a planned treatment period of 12 months. The primary research question was whether raloxifene compared with placebo improves cognitive function in women with AD. Participants were assigned to oral raloxifene hydrochloride given as two 60-mg tablets once daily or identical-appearing placebo. The dose was based on findings in the Multiple Outcomes of Raloxifene Evaluation trial, in which reductions in risk of cognitive impairment were seen with the 120 mg/d dose but not with the 60 mg/d dose.[4] Interventions took place on the campus of Indiana University School of Medicine and Kaiser Permanente Medical Center (Santa Rosa, CA). To facilitate recruitment, a third clinical site was established in 2008 on the campus of Southern Illinois University. Coordination and data management occurred at Stanford University. The study was monitored by an external data and safety monitoring board.

Standard protocol approvals, registrations, and patient consents. The study was approved by institutional review boards of Indiana University, Kaiser Permanente Division of Research, Southern Illinois University, and Stanford University. The protocol is registered at ClinicalTrials.gov (NCT00368459). Participants provided written informed consent, or assent with written consent of their next of kin or legally authorized representative.

Participants. Participants were postmenopausal women with dementia beginning at age 60 years or older, with probable AD[11] of mild to moderate severity, on a stable dose of a cholinesterase inhibitor. Women using raloxifene or menopausal hormone therapy were ineligible, but we did not exclude women using other drugs for osteoporosis treatment. Table e-1 on the *Neurology*® Web site at Neurology.org provides details of inclusion criteria.

Randomization. Trial eligibility was assessed at screening and confirmed at the coordination center. To ensure approximately equal allocation of women by dementia severity, assignment by the study epidemiologist to raloxifene or placebo in a 1:1 ratio used block randomization (concealed block size of 4) stratified by clinical site and dementia severity (mild severity, screening Mini-Mental State Examination[12] scores 20 to 26; moderate severity, scores 12 to 19). The trial epidemiologist prepared a randomization list using a random number generator (Proc Plan procedure, SAS version 9.2; SAS Institute, Cary, NC). The investigational drug pharmacist prepared and labeled study medications in bottles containing somewhat more than a 3-month supply (200 tablets). The randomization list included a unique, randomly assigned product identification number. Labeled bottles specifying "raloxifene 60 mg or placebo tablets" were maintained at clinical sites. On determination of trial eligibility and based on dementia severity stratum, the epidemiologist informed the site coordinator which product identification number was assigned to the participant. Participants, caregivers, staff, and investigators other than the epidemiologist and pharmacist were masked to group assignment.

Endpoints. In intent-to-treat analyses, the primary endpoint, compared between treatment groups, was change in the Alzheimer's Disease Assessment Scale, cognitive subscale[13] (ADAS-cog) assessed at baseline and 12 months. Planned secondary endpoints were change in global dementia rating, function, behavior, other cognitive test scores, caregiver burden, and caregiver distress. Outcomes were assessed at baseline and 6 and 12 months, and the ADAS-cog was also assessed at 3 and 9 months. Pill counts at each visit were used to assess adherence.

Change in global dementia rating was assessed with the Clinical Dementia Rating–Sum of Boxes score[14] obtained by a certified clinician using information from the participant and a knowledgeable informant. The standard Clinical Dementia Rating score was also recorded. Functional assessment used the Alzheimer's Disease Cooperative Study Activities of Daily Living Inventory.[15] Ratings on this 23-item, informant-based instrument are derived from basic and instrumental daily activities during the preceding 4 weeks. Behavior was assessed with the Neuropsychiatric Inventory[16] during a structured interview with the participant's caregiver informant, with items from 10 behavioral domains. Secondary cognitive outcomes[13,17–25] emphasized skills posited to be affected by a drug that interacts with estrogen receptors (verbal episodic memory, working memory and executive functions, and semantic memory). A global cognitive composite score was calculated as the weighted sum of component standardized scores weighted by the inverse intertest correlation matrix.[26] Caregiver burden was assessed by the 22-item self-report Zarit Caregiver Burden Interview,[27] and caregiver distress was rated with the 10-item Caregiver Distress Scale given as part of the Neuropsychiatric Interview.[16] At the study conclusion, informants indicated on a 5-point Likert scale whether they thought family members had received raloxifene or placebo.

Demographic and clinical variables. Demographic information, medical history, and medication use were obtained by structured questionnaires. Body mass index (kg/m^2) was calculated from measured height and weight.

Statistical analysis. For this pilot trial, we did not anticipate power to detect small, meaningful between-group differences over a 12-month treatment period but specified that we would report trends ($\alpha < 0.1$) as a guide to future studies. Planned recruitment of 72 participants was based on feasibility and anticipated resources. An intent-to-treat analysis was performed for all participants who completed baseline assessments, with missing data imputed from group mean changes. Standardized differences (effect sizes) were calculated from test means and SDs from baseline test data for the entire sample.

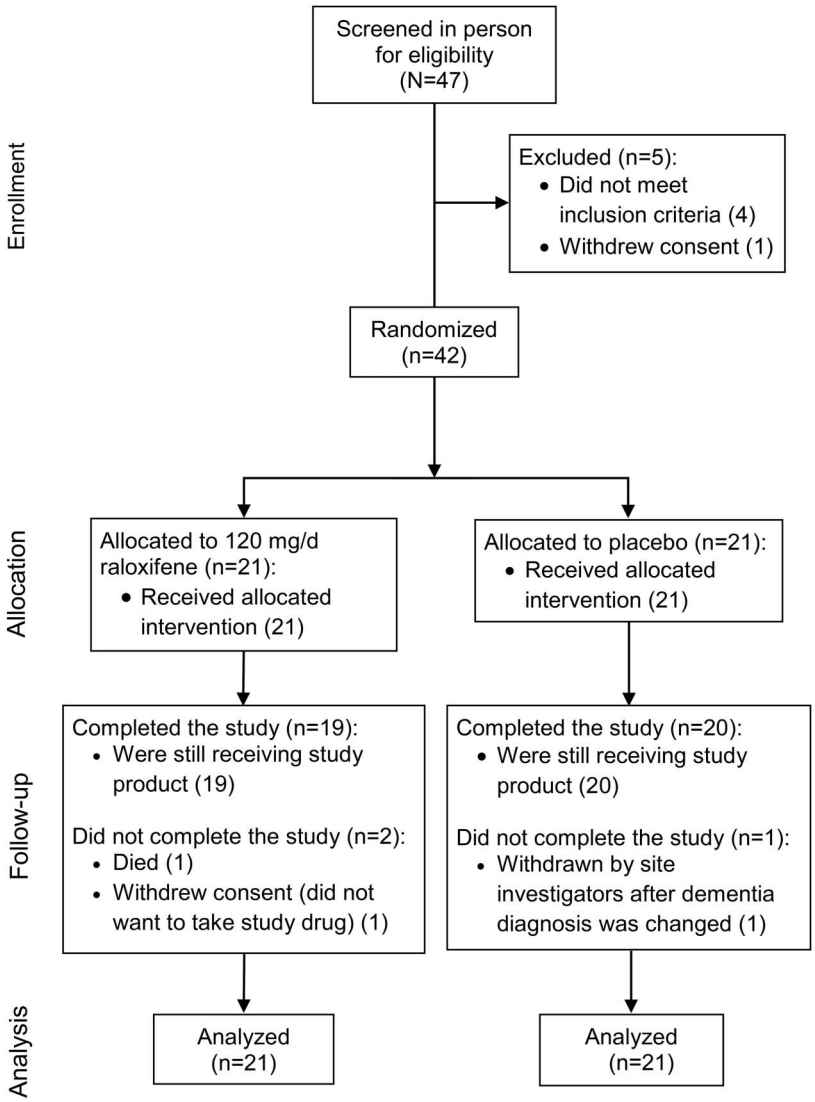
Initial analyses explored possible differences in baseline characteristics between treatment groups, analyzed by t test for continuous variables or χ^2 for categorical variables. Between-group change scores were analyzed by analysis of covariance with baseline scores, age, and education as covariates. An exploratory analysis of the ADAS-cog evaluated outcomes with repeated-measures general linear models according to a $5 \times 2 \times 3$ factorial design. The within-subjects factor was time (baseline, 3, 6, 9, and 12 months) and between-subjects factors were treatment (raloxifene, placebo)

and clinical site. We used SAS Proc Mixed to model the covariance structure and maintain cases with missing outcome data. We used polynomial contrasts to investigate the nature of the treatment-by-time interaction.

Classification of evidence. This interventional study provides Class I evidence that 12-month treatment with raloxifene 120 mg/d does not have a significant cognitive effect in women with late-onset AD of mild to moderate severity (mean between-groups standardized difference on the ADAS-cog of 0.03, 95% CI −0.39 to 0.44, 2-tailed $p = 0.89$). The study lacked the precision to exclude a small effect.

RESULTS Forty-seven women were assessed in person for eligibility, and 42 were randomly assigned to a treatment group (21 raloxifene, 21 placebo). Thirty-nine women provided baseline and 12-month outcome data. Three other women who provided baseline data were included in intent-to-treat analyses (figure). Women ranged in age from 68 to 84 years (table 1). Women in the raloxifene group were slightly older and slightly less well educated than women in the placebo group but were otherwise similar, and subsequent analyses adjusted for age and education. At baseline, ADAS-cog performance, dementia rating, function, and behavioral symptoms did not differ significantly, and treatment groups were otherwise similar (table 1; baseline neuropsychological test scores are given in table e-2). There was no effect of site on study outcomes, and analyses were pooled across sites.

Figure Flowchart of study enrollment and follow-up



Primary efficacy endpoint, ADAS-cog. At 12 months, women in both treatment groups showed modest declines in mean ADAS-cog performances (table 2). The 12-month decline in the placebo group was similar to that reported in other clinical trials involving patients with mild to moderate dementia.[28] The baseline ADAS-cog SD was 11.4 points, and the 12-month between-groups mean difference of 0.3 points represented a standardized difference (standardized effect size) of only 0.03 (95% CI −0.39 to 0.44). Effects were similar in subgroups defined by baseline severity (mild vs moderate dementia; interaction p value = 0.44, 2-tailed analysis of covariance).

Secondary endpoints. Dementia rating, function, and behavior declined in both treatment groups, without significant differences between groups (table 2). At 12 months, the standardized difference on the global cognitive composite score showed a medium effect size (>0.5 SD) in favor of raloxifene that, however, was not significant ($p = 0.38$, 2-tailed analysis of covariance) (table 2). For individual neuropsychological tests, there were no consistent between-group differences (table 2). We found a large (>0.8 SD), nominally significant effect favoring raloxifene for maze performance, and a medium, nonsignificant effect on one verbal memory task (word list delayed recall). We noted small (>0.2 SD), nonsignificant effect sizes favoring raloxifene on a related verbal memory task (word list immediate recall), on 2 writing measures, and on 2 measures of semantic memory. We noted a small, nonsignificant effect favoring placebo on a different semantic memory task. Caregiver burden and caregiver distress increased modestly over time in both treatment groups (table 2).

Exploratory analyses: Efficacy outcomes before 12 months. Women in the raloxifene group showed small improvements in the ADAS-cog at 3, 6, and 9 months; women in the placebo group declined at each time point (table 3). The standardized difference was nominally significant at 3 months, corresponding to a mean difference of 3.8 points on the ADAS-cog, but not at other time points. At 6 months, functional decline was greater in the raloxifene group. Declines

Table 1 Participant characteristics at baseline, by group

Characteristic[a]	Raloxifene group (n = 21)	Placebo group (n = 21)
Dementia severity,[b] n (%)		
Mild	13 (62)	12 (57)
Moderate	8 (38)	9 (43)
Age, y	77.7 (4.5)	74.1 (5.1)[a]
Education, y	13.2 (2.4)	14.0 (2.0)[a]
Race or ethnicity, n (%)		
White, non-Hispanic	20 (95)	21 (100)
Black, non-Hispanic	1 (5)	0 (0)
Surgical menopause, n (%)	9 (43)	8 (38)
Body mass index, kg/m^2	26.0 (5.0)	25.7 (4.7)
Systolic blood pressure, mm Hg	131 (17)	132 (24)
Diastolic blood pressure, mm Hg	66 (9)	68 (16)
ADAS-cog, error score[13] (70–0)	24.3 (11.0)	25.8 (12.0)
Mini-Mental State Examination[12] (0–30)	21.2 (4.9)	19.4 (6.2)
Clinical Dementia Rating[14] (3–0)	1.0 (0.5)	1.2 (0.6)
Clinical Dementia Rating–Sum of Boxes (18–0)	5.5 (3.0)	6.8 (3.1)
Activities of daily living[15] (0–78)	63.5 (12.4)	58.3 (14.8)
Neuropsychiatric Inventory[16] (120–0)	5.2 (6.6)	5.8 (7.8)
Caregiver burden interview score[27] (110–22)	38.6 (11.9)	43.7 (11.8)
Caregiver distress[16] (50–0)	4.0 (5.6)	2.8 (3.1)

Abbreviation: ADAS-cog = Alzheimer's Disease Assessment Scale, cognitive subscale.
Data are shown as n (%) or mean (SD). For baseline values of outcome measures, numbers in parentheses provide the range of possible scores. Higher values represent less impairment on activities of daily living but more impairment for the ADAS-cog, Clinical Dementia Rating and Clinical Dementia Rating–Sum of Boxes, the Neuropsychiatric Inventory, and caregiver burden and distress.
[a] For between-group comparisons—using Wilcoxon rank sum test, *t* test, or χ^2 test, as appropriate—all 2-tailed probabilities >0.1 except age ($p = 0.02$) and education ($p = 0.1$).
[b] Severity was defined by Mini-Mental State Examination scores of 20–26 (mild) and 12–19 (moderate).

in dementia rating and behavior—as well as changes on individual neuropsychological scores—did not differ significantly between treatment groups (table 4). Medium effect sizes (>0.5 SD) were noted on 2 tasks, one favoring raloxifene (a measure of semantic memory) and one favoring placebo (an executive functions task [Trail Making Test, part B]).

Other exploratory analyses. Mixed linear models predicted a nonsignificant difference in the annual rate of change in the ADAS-cog of 0.4 points (95% CI −3.0 to 3.8) favoring the raloxifene group. The effect of group is also very small if time is modeled as a quadratic function. The ADAS-cog may be relatively insensitive to cognitive change in patients with mild disease,[29] and the rate difference was more apparent among women with moderate dementia (n = 17), where the mean annual change in the placebo group was −7.6 (95% CI −3.0 to 12.2) points per year and the predicted difference was 3.3 (−2.7 to 9.4) points, favoring the raloxifene group. For women with mild dementia (n = 25), the mean annual change in the placebo group was −0.4 (−3.0 to 2.7) and the predicted between-groups difference was −1.6 (−5.6 to 2.4; interaction *p* value = 0.079).

Adherence. Based on pill count, the proportion of adherent women did not differ between treatment groups. Of the 39 study completers, 17 of 18 women in the raloxifene group (one with missing data) and 20 of 20 women in the placebo group took at least 80% of allocated study drug. The one nonadherent participant took 77% of her allotted study drug.

Safety. Three participants experienced serious adverse events, including one death. In the raloxifene group, one woman developed pneumonia and congestive heart failure, had an ischemic stroke, and died; another was diagnosed with colon cancer. A woman in the placebo group was hospitalized for hallucinations and agitation (adverse event details are in table e-3).

Adequacy of blinding. During the final visit, informants of 40 participants provided Likert scale responses regarding suspected treatment arm. Responses did not differ between treatment groups ($p = 0.75$, 2-tailed *t* test). Five informants from each group were "very certain" or "somewhat certain" that their family members had received raloxifene.

DISCUSSION Results from this randomized, double-blind, placebo-controlled pilot trial confirm the 12-month feasibility of high-dose (120 mg/d) raloxifene therapy in women with late-onset AD of mild to moderate severity. Ninety percent of raloxifene group participants contributed outcomes at 12 months, and all but one took at least 80% of the allocated treatment. Caregiver blinding was maintained. However, recruitment was more difficult than anticipated, and the study sample did not reflect the racial and ethnic diversity of the general population. One woman in the raloxifene group experienced a possibly related, serious adverse event: ischemic stroke followed by death. Raloxifene is linked to fatal stroke in women at high risk defined by the Framingham stroke score.[30] In future trials, high stroke risk could be considered exclusionary.

This pilot trial was not designed to identify small effects of raloxifene in the range provided by approved Alzheimer therapies such as donepezil or memantine. We found that raloxifene did not have a large or medium effect on the prespecified primary outcome. Based on the 12-month change of ADAS-cog performance compared between groups, the mean effect size was negligible, averaging only

Table 2 Twelve-month changes for primary and secondary outcomes

Outcome (range of potential scores)	Raloxifene group (n = 21)	Placebo group (n = 21)	Standardized difference[a] (95% CI)
Primary outcome			
ADAS-cog[13] (70–0)	−3.2	−3.5	0.03 (−0.39 to 0.44)
Noncognitive secondary outcomes			
Clinical Dementia Rating[14] (3–0)	−0.5	−0.3	−0.30 (−1.16 to 0.56)
Clinical Dementia Rating–Sum of Boxes (18–0)	−2.6	−2.0	−0.18 (−0.75 to 0.39)
Activities of daily living[15] (0–78)	−9.1	−4.5	−0.34 (−0.84 to 0.18)
Neuropsychiatric Inventory[16] (120–0)	−2.3	−2.5	0.02 (−0.65 to 0.70)
Cognitive secondary outcomes			
Word list learning,[13,17] immediate recall[b] (0–30)	0.3	−1.0	0.29 (−0.20 to 0.77)
Word list learning, delayed recall[b] (0–10)	0.5	−0.3	0.69 (−0.04 to 1.42)[c]
East Boston Memory Test,[18] immediate recall (0–12)	−1.5	−1.8	0.10 (−0.59 to 0.79)
East Boston Memory Test, delayed recall (0–12)	−0.1	−0.2	0.04 (−0.67 to 0.75)
Digit ordering,[19] span (0–6)	−0.3	−0.4	0.02 (−0.58 to 0.62)
Maze completion,[17] s (240–0)	−9.1	−81.0	0.89 (0.03 to 1.75)[d]
Number cancellation,[17] targets (0–40)	−1.3	−2.9	0.18 (−0.28 to 0.64)
Trail Making Test,[20] part A, s (300–0)	−30.8	−15.1	−0.16 (−0.64 to 0.31)
Trail Making Test, part B, s (300–0)	4.4	−2.2	0.09 (−0.64 to 0.81)
Category fluency,[21] animals (≥0)	−2.1	−2.9	0.16 (−0.35 to 0.66)
Boston Naming Test[22] (0–30)	−3.0	−1.2	−0.28 (−0.77 to 0.20)
Narrative writing[23] (0–11)	−0.2	−0.5	0.24 (−0.24 to 0.72)
Narrative writing, semantic density[24] (0–21)	−0.2	−1.3	0.36 (−0.09 to 0.82)
Semantic retrieval,[25] correct names (0–16)	−1.1	−1.9	0.24 (0.43 to 0.90)
Semantic retrieval, recognition (0–32)	−0.7	−2.0	0.32 (−0.34 to 0.98)
Visuoconstructive performance[e] (0–15)	−0.6	−1.0	0.11 (−0.29 to 0.50)
Mini-Mental State Examination[12] (0–30)	−2.5	−1.6	−0.17 (−0.65 to 0.31)
Cognitive composite score, standardized units[f]	−0.6	−1.1	0.52 (−0.62 to 1.66)
Caregiver secondary outcomes			
Caregiver burden interview score[27] (110–22)	−3.1	−2.0	−0.12 (−0.67 to 0.44)
Caregiver distress[16] (50–0)	−1.5	−1.0	−0.07 (−0.72 to 0.57)

Abbreviations: ADAS-cog = Alzheimer's Disease Assessment Scale, cognitive subscale; CI = confidence interval.
Group differences in mean test scores are adjusted for baseline score, age, and education. In all instances, a positive change indicates improved performance. Signs were changed for ADAS-cog, Clinical Dementia Rating, and Clinical Dementia Rating–Sum of Boxes, Neuropsychiatric Inventory, caregiver burden, and caregiver distress.
[a] Standardized difference (effect size), calculated as the mean between-group difference (raloxifene minus placebo) divided by the SD of the test; SDs for standardization used baseline test data for the entire sample.
[b] From the 10-item ADAS-cog[17] word list, with 3 immediate recall trials and one delayed recall trial.
[c] $p = 0.06$, 2-tailed analysis of covariance.
[d] $p = 0.04$, 2-tailed analysis of covariance.
[e] Extended scoring of ADAS-cog drawings[6] (circle, 2 points; diamond, 3 points; overlapping rectangles, 4 points; cube, 6 points).
[f] Calculated as the weighted average of standardized test scores for the ADAS-cog and cognitive secondary outcomes weighted by the inverse interest correlation matrix.[26]

three-tenths of a point, or 0.03 SD, on this test. Early cognitive benefit of raloxifene compared with placebo (0.33 SD at 3 months) was not sustained at later time points. Secondary analyses of dementia rating, activities of daily living, and behavior—as well as caregiver burden and distress—also failed to suggest meaningful benefit, although power to detect treatment effects was limited.

In secondary analyses based on mean change at 12 months, the global cognition composite measure showed a medium effect size (>0.5 SD) in favor of raloxifene that, however, was not significant; this

Table 3 Exploratory outcome: Change in ADAS-cog score between baseline and 3, 6, 9, and 12 months

Time point, mo	Mean change, raloxifene group	Mean change, placebo group	Between-group difference	Standardized difference (95% CI)	Nominal p value[a]
3	1.5	−2.3	3.8	0.33 (0.003 to 0.66)	0.048
6	0.7	−1.8	2.5	0.22 (−0.10 to 0.54)	0.18
9	1.1	−1.3	2.4	0.21 (−1.10 to 0.52)	0.17
12	−3.2	−3.5	0.3	0.03 (−0.39 to 0.44)	0.89

Abbreviations: ADAS-cog = Alzheimer's Disease Assessment Scale, cognitive subscale; CI = confidence interval.
Values represent mean change at each time point, adjusted for baseline score, age, and education. A positive mean change represents improved performance, compared with baseline. A positive between-group difference represents better performance by the raloxifene group.
[a] Two-tailed analysis of covariance, unadjusted for multiple comparisons.

Table 4 Exploratory outcomes: Six-month changes for cognitive and noncognitive outcomes

Outcome	Raloxifene group (n = 21)	Placebo group (n = 21)	Standardized difference (95% CI)
ADAS-cog	0.7	−1.8	0.22 (−0.10 to 0.54)
Noncognitive outcomes			
Clinical Dementia Rating	−0.2	−0.2	−0.08 (−0.74 to 0.57)
Clinical Dementia Rating-Sum of Boxes	−0.8	−1.0	0.07 (−0.44 to 0.58)
Activities of daily living	−6.9	−0.3	−0.48 (−0.82 to −0.15)[a]
Neuropsychiatric Inventory	−0.7	−3.1	0.35 (−0.35 to 1.04)
Caregiver outcomes			
Caregiver burden interview score	−2.0	−1.0	−0.08 (−0.65 to 0.49)
Caregiver distress	−0.2	−0.60	0.10 (−0.46 to 0.67)
Other cognitive outcomes			
Word list learning, immediate recall	0.2	0.1	0.03 (−0.27 to 0.33)
Word list learning, delayed recall	−0.04	−0.2	0.15 (−0.27 to 0.58)
East Boston Memory Test, immediate recall	−0.4	−0.5	0.02 (−0.59 to 0.63)
East Boston Memory Test, delayed recall	0.4	−0.5	0.40 (−0.23 to 1.02)
Digit ordering, span	−0.1	−0.05	−0.07 (−0.51 to 0.36)
Maze completion, s	−24.0	−15.9	−0.10 (−0.69 to 0.49)
Number cancellation, no. of targets	−0.1	−1.4	0.14 (−0.35 to 0.63)
Trail Making Test, part A, s	−16.3	−7.6	−0.09 (−0.51 to 0.33)
Trail Making Test, part B, s	−28.9	10.1	−0.51 (−1.11 to 0.08)
Category fluency, animals	−1.6	−1.4	−0.05 (−0.42 to 0.32)
Boston Naming Test	−1.7	−0.4	−0.22 (−0.56 to 0.13)
Narrative writing	0.01	−0.01	0.02 (−0.46 to 0.50)
Narrative writing, semantic density	0.3	−0.2	0.15 (−0.29 to 0.60)
Semantic retrieval, correct names	0.03	−1.6	0.44 (−0.08 to 0.95)
Semantic retrieval, recognition	−0.7	−2.8	0.54 (−0.32 to 1.40)
Visuoconstructive performance	−0.6	−1.0	0.11 (−0.29 to 0.50)
Mini-Mental State Examination	−0.8	−0.6	−0.05 (−0.44 to 0.35)
Cognitive composite score, standardized units	−0.5	−0.6	0.16 (−0.74 to 1.07)

Abbreviations: ADAS-cog = Alzheimer's Disease Assessment Scale, cognitive subscale; CI = confidence interval.
Group differences in mean test scores are adjusted for baseline score, age, and education. In all instances, a positive change indicates improved performance. Signs were changed for ADAS-cog, Clinical Dementia Rating and Clinical Dementia Rating-Sum of Boxes, Neuropsychiatric Inventory, caregiver burden, and caregiver distress.
[a] $p = 0.006$, 2-tailed analysis of covariance.

difference was smaller at 6 months. Nonsignificant differences favoring raloxifene on some verbal memory tasks could have been observed on the basis of chance. However, raloxifene is sometimes reported to benefit verbal memory in older postmenopausal women without dementia,[5,8] and future studies might examine raloxifene effects within this cognitive domain more closely.

These results provide information to guide consideration and design of future trials. The essentially null effect of raloxifene on the primary outcome implies a low likelihood of positive results but does not exclude the possibility of modest cognitive benefit or harm. The upper tail of the 95% CI represents a small ($0.2 \leq SD < 0.5$) effect but corresponds to about 5 points on the ADAS-cog, a difference that—despite weak evidence for validity[31]—falls within a commonly accepted range of clinical relevance. The ADAS-cog is a common endpoint in Alzheimer treatment trials, but it is possible that other neuropsychological outcomes could be more sensitive to any cognitive effect of raloxifene.

We conclude that 12 months' treatment with raloxifene 120 mg/d compared with placebo has no more than a small effect on the ADAS-cog in women with late-onset AD of mild to moderate severity. This conclusion may not pertain to short-term (e.g., 3-month) effects of raloxifene, where we found a nominally significant difference between groups on the ADAS-cog; to effects that might require more than 12 months to emerge; or to selective effects within specific cognitive domains. Finally, results do not generalize to populations or outcomes not studied in this trial, including raloxifene effects on cognitive outcomes in women without dementia, raloxifene effects on AD risk, or effects of other SERMs on cognitive outcomes.

AUTHOR CONTRIBUTIONS

V.W. Henderson conceptualized and designed the trial. T. Ala, K.L. Sainani, A.L. Bernstein, B.S. Stephenson, and M.R. Farlow collected the data. K.L. Sainani performed the statistical analyses. V.W. Henderson and K.L. Sainani interpreted the data. V.W. Henderson wrote the manuscript. All authors critically revised and approved the final manuscript.

ACKNOWLEDGMENT

Alzheimer's Disease Assessment Scale, cognitive subscale forms were used with permission from the NIA Alzheimer's Disease Cooperative Study (AG10483). The Neuropsychiatric Inventory was used with permission of Dr. Jeffrey Cummings. The authors are grateful to the study participants and their caregivers, and thank other contributors to this study: Katherine Cavanaugh (Indiana University, psychometrician); Ann M. Hake, MD (Indiana University, study neurologist); Sheryl E. Lynch, LPN (Indiana University, site coordinator); Brandy R. Matthews, MD (Indiana University, study neurologist); Karen Turchi, MSW (Indiana University, psychometrician); Rodger J. Elble, MD, PhD (Southern Illinois University, study facilitator); Barbara C. Lokaitis, BA (Southern Illinois University, site coordinator); Ronald F. Zec, PhD (Southern Illinois University, neuropsychologist); Terry Hess, MA (Kaiser Permanente Medical Center, site coordinator); Susan Kirk, RN (Kaiser Permanente Medical Center, site coordinator); Christine M. Naber, PhD (Kaiser Permanente Medical Center, neuropsychologist); Ann Werlin, RN (Kaiser Permanente Medical Center, site coordinator); Narinder Bolaria, MS (Stanford University, study coordinator); Martha E. Hamilton, PharmD (Stanford University, investigational drug pharmacist); Cheryl Langford (Stanford University, study coordinator); Rita A. Popat, PhD (Stanford University, study epidemiologist); Douglas R. Galasko, MD (University of California, San Diego, Data and Safety Monitoring Board chair); Jody Corey-Bloom, MD, PhD (University of California, San Diego, Data and Safety Monitoring Board member); and Donna Shoupe, MD (University of Southern California, Data and Safety Monitoring Board member).

STUDY FUNDING

Supported by NIH (R01-AG023038). Eli Lilly and Company (Indianapolis, IN) provided study drug and placebo without charge or restriction.

DISCLOSURE

V. Henderson receives research support from the NIH. T. Ala receives research support from Forum Pharmaceuticals, Merck & Co., Navidea Biopharmaceuticals, Toyama Chemical, and Transition Therapeutics. K. Sainaini receives payment as the statistical editor for *PM&R*. A. Bernstein receives research support from Accera, Biogen, Eli Lilly and Co., Genentech, and Lundbeck. B. Stephenson and A. Rosen report no disclosures relevant to the manuscript. M. Farlow receives research support from Accera, Biogen, Boehringer Ingelheim, Chase Pharmaceuticals, Eisai, Eli Lilly and Co., Genentech, Lundbeck, and Roche; is a paid speaker for Eisai, Forest, Novartis, Eli Lilly and Co., and Pfizer; and serves as a paid consultant for Accera, Alltech, Avanir, Biogen, Eisai Medical Research, Eli Lilly and Co., FORUM Pharmaceuticals, Grifols, Helicon, INC Research, Lundbeck, MedAvante, Medivation Inc., Merck & Co. Inc., Neurotrope Biosciences, Novartis, Pfizer, Prana Biotech, QR Pharma, Roivant Sciences, Roche, Sanofi-Aventis, Schering-Plough, and UCB Pharma. Go to Neurology.org for full disclosures.

Received April 15, 2015. Accepted in final form August 6, 2015.

REFERENCES

1. Barnes LL, Wilson RS, Bienias JL, Schneider JA, Evans DA, Bennett DA. Sex differences in the clinical manifestations of Alzheimer disease pathology. Arch Gen Psychiatry 2005;62:685–691.
2. Henderson VW, Buckwalter JG. Cognitive deficits of men and women with Alzheimer's disease. Neurology 1994;44:90–96.
3. Henderson VW. Alzheimer's disease: review of hormone therapy trials and implications for treatment and prevention after menopause. J Steroid Biochem Mol Biol 2014;142:99–106.
4. Yaffe K, Krueger K, Cummings SR, et al. Effect of raloxifene on the prevention of dementia and cognitive impairment in older women: the Multiple Outcomes of Raloxifene Evaluation (MORE) randomized trial. Am J Psychiatry 2005;162:683–690.
5. Yaffe K, Krueger K, Sarkar S, et al. Cognitive function in postmenopausal women treated with raloxifene. N Engl J Med 2001;344:1207–1213.
6. Haskell SG, Richardson ED. The effect of raloxifene on cognitive function in postmenopausal women: a randomized clinical trial. Conn Med 2004;68:355–358.
7. Espeland MA, Shumaker SA, Limacher M, et al. Relative effects of tamoxifen, raloxifene, and conjugated equine estrogens on cognition. J Womens Health 2010;19:371–379.
8. Jacobsen DE, Melis RJ, Verhaar HJ, Olde Rikkert MG. Raloxifene and tibolone in elderly women: a randomized, double-blind, double-dummy, placebo-controlled trial. J Am Med Dir Assoc 2012;13:189.e1–189.e7.

9. Nickelsen T, Lufkin EG, Riggs BL, Cox DA, Crook TH. Raloxifene hydrochloride, a selective estrogen receptor modulator: safety assessment of effects on cognitive function and mood in postmenopausal women. Psychoneuroendocrinology 1999;24:115–128.
10. Buckwalter JG, Geiger AM, Parsons TD, Handler J, Howes J, Lehmer RR. Cognitive effects of short-term use of raloxifene: a randomized clinical trial. Int J Neurosci 2007;117:1579–1590.
11. McKhann G, Drachman D, Folstein MF, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. Neurology 1984;34:939–944.
12. Folstein MF, Folstein SE, McHugh PR. "Mini-Mental State": a practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975;12:189–198.
13. Rosen WG, Mohs RC, Davis KL. A new rating scale for Alzheimer's disease. Am J Psychiatry 1984;141:1356–1364.
14. Morris JC. The Clinical Dementia Rating (CDR): current version and scoring rules. Neurology 1993;43:2412–2414.
15. Galasko D, Bennett D, Sano S, et al. An inventory to assess activities of daily living for clinical trials in Alzheimer's disease: the Alzheimer's Disease Cooperative Study. Alzheimer Dis Assoc Disord 1997;11(suppl 2):S33–S39.
16. Cummings JL, Mega M, Gray K, Rosenberg-Thompson S, Carusi DA, Gornbein J. The Neuropsychiatric Inventory: comprehensive assessment of psychopathology. Neurology 1994;44:2308–2314.
17. Mohs RC, Knopman D, Petersen RC, et al. Development of cognitive instruments for use in clinical trials of antidementia drugs: additions to the Alzheimer's Disease Assessment Scale that broaden its scope. The Alzheimer's Disease Cooperative Study. Alzheimer Dis Assoc Disord 1997;11(suppl 2):S13–S21.
18. Scherr PA, Albert MS, Funkenstein HH, et al. Correlates of cognitive function in an elderly community population. Am J Epidemiol 1988;128:1084–1101.
19. MacDonald MC, Almor A, Henderson VW, Kempler D, Andersen ES. Assessing working memory and language comprehension in Alzheimer's disease. Brain Lang 2001;78:17–42.
20. Reitan RM. Validity of the Trail Making Test as an indicator of organic brain damage. Percept Mot Skills 1958;8:271–276.
21. Troyer AK. Normative data for clustering and switching on verbal fluency tasks. J Clin Exp Neuropsychol 2000;22:370–378.
22. Mack WJ, Freed DM, Williams BW, Henderson VW. Boston Naming Test: shortened versions for use in Alzheimer's disease. J Gerontol 1992;47:P154–P158.
23. Goodglass H, Kaplan E, Barresi B. The Assessment of Aphasia and Related Disorders, 3rd ed. Philadelphia: Lippincott Williams & Wilkins; 2001.
24. Henderson VW, Buckwalter JG, Sobel E, Freed DM, Diz MM. The agraphia of Alzheimer's disease. Neurology 1992;42:776–784.
25. Kraut MA, Cherry B, Pitcock JA, Vestal L, Henderson VW, Hart JJ. The Semantic Object Retrieval Test (SORT) in normal aging and Alzheimer disease. Cogn Behav Neurol 2006;19:177–184.
26. Henderson VW, St John JA, Hodis HN, et al. Long-term soy isoflavone supplementation and cognition in women: a randomized, controlled trial. Neurology 2012;78:1841–1848.
27. Zarit SH, Reever KE, Bach-Peterson J. Relatives of the impaired elderly: correlates of feelings of burden. Gerontologist 1980;20:649–655.
28. Gomeni R, Simeoni M, Zvartau-Hind M, Irizarry MC, Austin D, Gold M. Modeling Alzheimer's disease progression using the disease system analysis approach. Alzheimers Dement 2012;8:39–50.
29. Wattmo C, Wallin AK, Minthon L. Progression of mild Alzheimer's disease: knowledge and prediction models required for future treatment strategies. Alzheimers Res Ther 2013;5:44.
30. Barrett-Connor E, Cox DA, Song J, Mitlak B, Mosca L, Grady D. Raloxifene and risk for stroke based on the Framingham Stroke Risk Score. Am J Med 2009;122:754–761.
31. Oremus M. Does the evidence say a 4-point change in ADAS-cog score is clinically significant? Alzheimers Dement 2014;10:416–417.